

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-258038

(43)Date of publication of application : 21.09.2001

(51)Int.Cl. H04N 7/32

// H04N 5/92

(21)Application number : 2001-024529

(71)Applicant : CANON INC

(22)Date of filing : 31.01.2001

(72)Inventor : JIN U

LENNON ALISON JOAN

(30)Priority

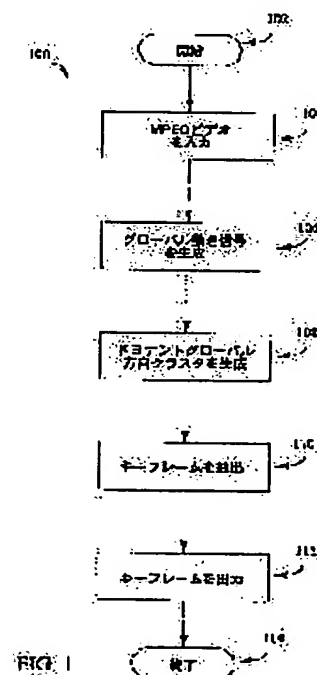
Priority number : 2000 PQ5352 Priority date : 31.01.2000 Priority country : AU

(54) METHOD FOR EXTRACTING KEY FRAME OF VIDEO SEQUENCE, KEY FRAME EXTRACTING DEVICE FOR VIDEO SEQUENCE, STORAGE MEDIUM AND PROGRAM SOFTWARE

(57)Abstract:

PROBLEM TO BE SOLVED: To substantially eliminate one or more defects of an existing configuration or to at least improve the defect.

SOLUTION: The method for extracting a key frame from a video sequence consisting of MPEG-compressed video data having a block motion vector, includes a step (104) where the MPEG-compressed video data are received, a step (106) where the received data are partially decompressed to acquire a block motion vector, and the block motion vector is converted into a forward block motion vector to generate a global motion signal, a step (108) where a dominant global direction cluster is generated, and a step (110) where the generated dominant global direction cluster is used to select a potential key frame of the video sequence, and the selected, potential key frame is decompressed to extract the key frame.



(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号  
特開2001-258038  
(P2001-258038A)

(43)公開日 平成13年9月21日(2001.9.21)

(51)Int.Cl. <sup>7</sup>	識別記号	F I	テーマコード(参考)
H 0 4 N 7/32		H 0 4 N 7/137	Z
// H 0 4 N 5/92		5/92	H

審査請求 未請求 請求項の数27 O L (全 16 頁)

(21)出願番号 特願2001-24529(P2001-24529)  
(22)出願日 平成13年1月31日(2001.1.31)  
(31)優先権主張番号 P Q 5 3 5 2  
(32)優先日 平成12年1月31日(2000.1.31)  
(33)優先権主張国 オーストラリア (AU)

(71)出願人 000001007  
キヤノン株式会社  
東京都大田区下丸子3丁目30番2号  
(72)発明者 ジン ウ  
オーストラリア国 2113 ニュー サウス  
ウェールズ州, ノース ライド, トーマ  
ス ホルト ドライブ 1 キヤノン イ  
ンフォメーション システムズ リサーチ  
オーストラリア プロプライエタリー  
リミテッド 内  
(74)代理人 100076428  
弁理士 大塚 康德 (外1名)

最終頁に続く

(54)【発明の名称】 ビデオシーケンスのキーフレーム抽出方法、ビデオシーケンスのキーフレーム抽出装置、記憶媒体及びプログラムソフトウェア

(57)【要約】

【課題】 既存の構成の1つ又は複数の欠点を実質的に解消するか、少なくとも改善する。

【解決手段】 ブロック動きベクトルを持つMPEG圧縮ビデオデータを有するビデオシーケンスからキーフレームを抽出する方法。この方法は最初、MPEG圧縮ビデオデータを入力し(104)、部分的に解凍してブロック動きベクトルを取得し、そのブロック動きベクトルを順方向ブロック動きベクトルに変換してグローバル動き信号を生成する(106)。そして、ドミナントグローバル方向クラスタを生成し(108)、生成されたドミナントグローバル方向クラスタを使用してビデオシーケンスの潜在的なキーフレームを選択し、選択された潜在的なキーフレームを解凍し、キーフレームを抽出する(110)。

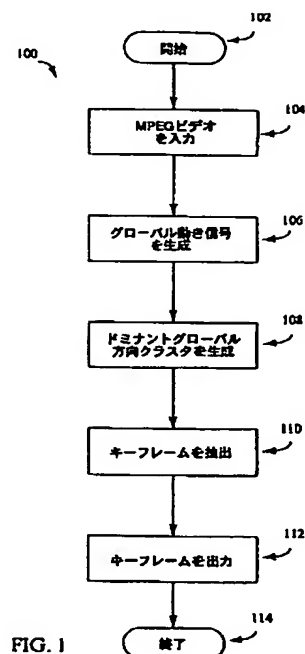


FIG. 1

## 【特許請求の範囲】

【請求項 1】 動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスからキーフレームを抽出する方法であって、

前記動きベクトルに基づいてグローバル動き信号を生成する工程と、

生成されたグローバル動き信号に基づいてドミナントグローバル方向クラスタを生成する工程と、

生成されたドミナントグローバル方向クラスタを使用してキーフレームを選択する工程と、

キーフレームを抽出するために前記選択されたキーフレームを解凍する工程とを有することを特徴とする方法。

【請求項 2】 前記選択する工程は、前記生成されたドミナントグローバル方向クラスタ及び定義済み規則のセットを使用してキーフレームを選択することを特徴とする請求項 1 に記載の方法。

【請求項 3】 前記選択する工程は、前記生成されたドミナントグローバル方向クラスタ及び第 1 の定義済み規則のセットを使用してビデオシーケンスの潜在的なキーフレームを選択し、

第 2 の定義済み規則のセットを使用して前記選択された潜在的なキーフレームから冗長なキーフレームを削除して選択されたキーフレームを得ることを特徴とする請求項 1 に記載の方法。

【請求項 4】 前記選択する工程は、前記生成されたドミナントグローバル方向クラスタ及び第 1 の定義済み規則のセットを使用してビデオシーケンスの潜在的なキーフレームを選択し、

第 2 の定義済み規則のセットを使用して前記選択された潜在的なキーフレームから冗長なキーフレームを削除して前記選択された潜在的なキーフレームのセットを得て、

色ヒストグラム技法を使用して前記選択された潜在的なキーフレームのセットから類似したキーフレーム及び／又は繰り返されるキーフレームを削除して前記選択されたキーフレームを得ることを特徴とする請求項 1 に記載の方法。

【請求項 5】 前記グローバル動き信号を生成する工程は、パングローバル動き信号、ズームグローバル動き信号、チルトグローバル動き信号を生成することを特徴とする請求項 1 に記載の方法。

【請求項 6】 前記ドミナントグローバル方向クラスタは、左パン動きクラスタ、右パン動きクラスタ、上チルト動きクラスタ、下チルト動きクラスタ、ズームイン動きクラスタ、ズームアウト動きクラスタ、グローバル静止動きクラスタの 1 つ又は複数を有することを特徴とする請求項 1 に記載の方法。

【請求項 7】 前記ドミナントグローバル方向クラスタを生成する工程は、

前記生成されたグローバル動き信号から離散グローバル

動き信号を生成し、

生成された離散グローバル動き信号からノイズを除去し、

前記ノイズが低減された離散グローバル動き信号に基づいてドミナントグローバル方向クラスタを生成することを特徴とする請求項 1 に記載の方法。

【請求項 8】 ブロック動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスの表現を生成する方法であって、

10 前記ブロック動きベクトルを得るために前記圧縮ビデオデータを解凍する工程と、

前記ブロック動きベクトルを順方向ブロック動きベクトルに変換する工程と、

前記順方向ブロック動きベクトルに基づいてグローバル動き信号を生成する工程と、

生成されたグローバル動き信号に基づいてドミナントグローバル方向クラスタを生成する工程と、

20 生成されたドミナントグローバル方向クラスタ及び定義済み規則のセットを使用してビデオシーケンスの潜在的なキーフレームを選択する工程と、

選択された潜在的なキーフレームの冗長なキーフレームを削除して残りの選択されたキーフレームを得る工程と、

前記ビデオシーケンスの表現を得るために前記残りの選択されたキーフレームを解凍する工程とを有することを特徴とする方法。

【請求項 9】 前記グローバル動き信号を生成する工程は、パングローバル動き信号、ズームグローバル動き信号、チルトグローバル動き信号を生成することを特徴とする請求項 8 に記載の方法。

30 【請求項 10】 前記ドミナントグローバル方向クラスタは、左パン動きクラスタ、右パン動きクラスタ、上チルト動きクラスタ、下チルト動きクラスタ、ズームイン動きクラスタ、ズームアウト動きクラスタ、グローバル静止動きクラスタの 1 つ又は複数を有することを特徴とする請求項 8 に記載の方法。

【請求項 11】 前記ドミナントグローバル方向クラスタを生成する工程は、

40 前記生成されたグローバル動き信号から離散グローバル動き信号を生成し、

生成された離散グローバル動き信号からノイズを除去し、

前記ノイズが低減された離散グローバル動き信号に基づいてドミナントグローバル方向クラスタを生成することを特徴とする請求項 8 に記載の方法。

【請求項 12】 各ビデオクリップがブロック動きベクトルを持つ M P E G 圧縮ビデオデータを有する、1 つ又は複数のビデオクリップからキーフレームを抽出する方法であって、

50 前記ブロック動きベクトルを得るために M P E G 圧縮ビ

デオデータを部分的に解凍する工程と、  
 前記ブロック動きベクトルを順方向ブロック動きベクトルに変換する工程と、  
 前記順方向ブロック動きベクトルに基づいてパングローバル動き信号、ズームグローバル動き信号、チルトグローバル動き信号を生成する工程と、  
 生成されたパングローバル動き信号、チルトグローバル動き信号、ズームグローバル動き信号に基づいて左パン動きクラスタ、右パン動きクラスタ、上チルト動きクラスタ、下チルト動きクラスタ、ズームイン動きクラスタ、ズームアウト動きクラスタ、グローバル静止動きクラスタの1つ又は複数を含むドミナントグローバル方向クラスタを生成する工程と、  
 生成されたドミナントグローバル方向クラスタ及び定義済み規則のセットを使用して各ビデオクリップの潜在的なキーフレームを選択する工程と、  
 定義済み発見的規則を使用して前記選択された潜在的なキーフレームから冗長なキーフレームを削除して残りの選択されたキーフレームの第1のセットを得る工程と、  
 色ヒストグラム技法を使用して前記残りの選択されたキーフレームの第1のセットから類似したキーフレーム及び／又は繰り返されるキーフレームを削除して残りの選択されたキーフレームの第2のセットを得る工程と、  
 キーフレームを抽出するために前記残りの選択されたキーフレームの第2のセットを解凍する工程とを有することを特徴とする方法。  
 【請求項13】 前記ドミナントグローバル方向クラスタを生成する工程は、  
 前記グローバル動き信号から離散グローバル動き信号を生成し、  
 生成された離散グローバル動き信号からノイズを除去し、  
 前記ノイズが低減された離散グローバル動き信号に基づいてドミナントグローバル方向クラスタを生成することを特徴とする請求項12に記載の方法。  
 【請求項14】 動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスからキーフレームを抽出する装置であって、  
 前記動きベクトルに基づいてグローバル動き信号を生成する手段と、  
 生成されたグローバル動き信号に基づいてドミナントグローバル方向クラスタを生成する手段と、  
 生成されたドミナントグローバル方向クラスタを使用してキーフレームを選択する手段と、  
 キーフレームを抽出するために前記選択されたキーフレームを解凍する手段とを有することを特徴とする装置。  
 【請求項15】 ブロック動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスの表現を生成する装置であって、  
 前記ブロック動きベクトルを得るために前記圧縮ビデオ

データを解凍する手段と、  
 前記ブロック動きベクトルを順方向ブロック動きベクトルに変換する手段と、  
 前記順方向ブロック動きベクトルに基づいてグローバル動き信号を生成する手段と、  
 生成されたグローバル動き信号に基づいてドミナントグローバル方向クラスタを生成する手段と、  
 生成されたドミナントグローバル方向クラスタ及び定義済み規則のセットを使用してビデオシーケンスの潜在的なキーフレームを選択する手段と、  
 選択された潜在的なキーフレームの冗長なキーフレームを削除して残りの選択されたキーフレームを得る手段と、  
 前記ビデオシーケンスの表現を得るために前記残りの選択されたキーフレームを解凍する手段とを有することを特徴とする装置。  
 【請求項16】 各ビデオクリップがブロック動きベクトルを持つMPEG圧縮ビデオデータを有する、1つ又は複数のビデオクリップからキーフレームを抽出する装置であって、  
 前記ブロック動きベクトルを得るためにMPEG圧縮ビデオデータを部分的に解凍する手段と、  
 前記ブロック動きベクトルを順方向ブロック動きベクトルに変換する手段と、  
 前記順方向ブロック動きベクトルに基づいてパングローバル動き信号、ズームグローバル動き信号、チルトグローバル動き信号を生成する手段と、  
 生成されたパングローバル動き信号、チルトグローバル動き信号、ズームグローバル動き信号に基づいて左パン動きクラスタ、右パン動きクラスタ、上チルト動きクラスタ、下チルト動きクラスタ、ズームイン動きクラスタ、ズームアウト動きクラスタ、グローバル静止動きクラスタの1つ又は複数を含むドミナントグローバル方向クラスタを生成する手段と、  
 生成されたドミナントグローバル方向クラスタ及び定義済み規則のセットを使用して各ビデオクリップの潜在的なキーフレームを選択する手段と、  
 定義済み発見的規則を使用して前記選択された潜在的なキーフレームから冗長なキーフレームを削除して残りの選択されたキーフレームの第1のセットを得る手段と、  
 色ヒストグラム技法を使用して前記残りの選択されたキーフレームの第1のセットから類似したキーフレーム及び／又は繰り返されるキーフレームを削除して残りの選択されたキーフレームの第2のセットを得る手段と、  
 キーフレームを抽出するために前記残りの選択されたキーフレームの第2のセットを解凍する手段とを有することを特徴とする装置。  
 【請求項17】 動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスからキーフレームを抽出するコンピュータプログラムを有するコンピュータ読み取り

可能な媒体であって、前記コンピュータプログラムは、前記動きベクトルに基づいてグローバル動き信号を生成するコードと、

生成されたグローバル動き信号に基づいてドミナントグローバル方向クラスタを生成するコードと、

生成されたドミナントグローバル方向クラスタを使用してキーフレームを選択するコードと、

キーフレームを抽出するために前記選択されたキーフレームを解凍するコードとを有することを特徴とするコンピュータ読み取り可能な媒体。

【請求項18】 ブロック動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスの表現を生成するコンピュータプログラムを有するコンピュータ読み取り可能な媒体であって、前記コンピュータプログラムは、前記ブロック動きベクトルを得るために前記圧縮ビデオデータを解凍するコードと、

前記ブロック動きベクトルを順方向ブロック動きベクトルに変換するコードと、

前記順方向ブロック動きベクトルに基づいてグローバル動き信号を生成するコードと、

生成されたグローバル動き信号に基づいてドミナントグローバル方向クラスタを生成するコードと、

生成されたドミナントグローバル方向クラスタ及び定義済み規則のセットを使用してビデオシーケンスの潜在的なキーフレームを選択するコードと、

選択された潜在的なキーフレームの冗長なキーフレームを削除して残りの選択されたキーフレームを得るコードと、

前記ビデオシーケンスの表現を得るために前記残りの選択されたキーフレームを解凍するコードとを有することを特徴とするコンピュータ読み取り可能な媒体。

【請求項19】 各ビデオクリップがブロック動きベクトルを持つMPEG圧縮ビデオデータを有する、1つ又は複数のビデオクリップからキーフレームを抽出するコンピュータプログラムを有するコンピュータ読み取り可能な媒体であって、前記コンピュータプログラムは、前記ブロック動きベクトルを得るためにMPEG圧縮ビデオデータを部分的に解凍するコードと、

前記ブロック動きベクトルを順方向ブロック動きベクトルに変換するコードと、

前記順方向ブロック動きベクトルに基づいてパングローバル動き信号、ズームグローバル動き信号、チルトグローバル動き信号を生成するコードと、

生成されたパングローバル動き信号、チルトグローバル動き信号、ズームグローバル動き信号に基づいて左パン動きクラスタ、右パン動きクラスタ、上チルト動きクラスタ、下チルト動きクラスタ、ズームイン動きクラスタ、ズームアウト動きクラスタ、グローバル静止動きクラスタの1つ又は複数を含むドミナントグローバル方向クラスタを生成するコードと、

生成されたドミナントグローバル方向クラスタ及び定義済み規則のセットを使用して各ビデオクリップの潜在的なキーフレームを選択するコードと、

定義済み発見的規則を使用して前記選択された潜在的なキーフレームから冗長なキーフレームを削除して残りの選択されたキーフレームの第1のセットを得るコードと、

色ヒストグラム技法を使用して前記残りの選択されたキーフレームの第1のセットから類似したキーフレーム及び／又は繰り返されるキーフレームを削除して残りの選択されたキーフレームの第2のセットを得るコードと、

キーフレームを抽出するために前記残りの選択されたキーフレームの第2のセットを解凍するコードとを有することを特徴とするコンピュータ読み取り可能な媒体。

【請求項20】 動きベクトルを持つ圧縮符号化されたビデオデータを含むビデオシーケンスからキーフレームを抽出する方法であって、

前記動きベクトルに基づいてグローバル動き信号を生成する工程と、

20 前記生成されたグローバル動き信号に基づいてキーフレームを選択する工程と、

前記選択されたキーフレームの圧縮符号化されたビデオデータを解凍する工程とを有することを特徴とする方法。

【請求項21】 前記選択する工程は、前記生成されたグローバル動き信号に基づいてフレーム分類を行なう工程を含み、前記分類結果に応じてキーフレームを選択することを特徴とする請求項20に記載の方法。

30 【請求項22】 前記解凍されたビデオデータを表示する工程を有することを特徴とする請求項20又は21に記載の方法。

【請求項23】 前記圧縮符号化されたビデオデータは、イントラ符号化処理とインター符号化処理とを適応的に用いて符号化されたデータであることを特徴とする請求項20乃至22のいずれか1項に記載の方法。

【請求項24】 前記圧縮符号化されたビデオデータは、MPEG(MovingPicture Coding Experts Group)符号化方式により符号化されたデータであることを特徴とする請求項23に記載の方法。

40 【請求項25】 請求項20乃至24のいずれか1項に記載のキーフレームを抽出する方法を実行するためのコンピュータ読み取り可能なプログラムコードを記憶した記憶媒体。

【請求項26】 請求項20乃至24のいずれか1項に記載のキーフレームを抽出する方法をコンピュータに実行するためのプログラムソフトウェア。

【請求項27】 動きベクトルを持つ圧縮符号化されたビデオデータを含むビデオシーケンスからキーフレームを抽出する装置であって、

50 前記動きベクトルに基づいてグローバル動き信号を生成

する手段と、  
前記生成されたグローバル動き信号に基づいてキーフレームを選択する手段と、  
前記選択されたキーフレームの圧縮符号化されたビデオデータを解凍する手段とを有することを特徴とする装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、全般的にビデオシーケンスからキーフレームを抽出することに関する。特に、本発明はビデオシーケンスからキーフレームを抽出する方法及び装置と、ビデオシーケンスの表現を生成する方法及び装置に関する。また、本発明は、このような方法を実現するコンピュータプログラムを有するコンピュータ読み取り可能な媒体にも関する。更に、本発明は、このような方法により作成されるビデオサマリにも関する。

【0002】

【従来の技術】近年、ビデオカメラの人気の益々高まってきた。カメラユーザは、一般に、各ビデオカセット又はその他の媒体上に1つ又は複数のビデオクリップを格納している。ビデオデータの増加に伴い、ユーザは自分のビデオデータを構成及び管理することが必要になっている。

【0003】ビデオデータを構成及び管理する1つの基本的な方法は、ビデオの特定部分をアクセスするためのキーワード・ベースの探索及び早送り／巻戻しブラウジングを必要とする。しかしながら、キーワード・ベースのデータ検索システムでは、ビデオデータの内容を厳密かつ一意に表現することはできない。早送り／巻戻し操作は極めて低速でかつ非効率的である。

【0004】ビデオクリップの特定部分にアクセスするための他の一般的な方法は、ビデオシーケンスから抽出される代表的なフレームとしてキーフレームを用いる。キーフレームの抽出は高速ビデオブラウジングを可能とし、またビデオ内容の要約化及び視覚化の強力なツールを提供する。

【0005】しかしながら、一定時間の求めによりフレームの抽出に基づくビデオの要約化及び視覚化は、ショット情報もフレームの類似性も使用していない。短い重要なショットの場合は代表的なフレームがない可能性があり、長いショットの場合は類似した内容を有する複数のフレームがある可能性がある。

【0006】ビデオサマリを作成するための他の一般的な方法は、カット／変化検出を使用して映画のショットに関する代表的なキーフレームを選択することである。代表的なフレームを選択するための典型的な手法は、カットポイントをキーフレームとして使用することである。そして、キーフレームはサマリとして使用される。通常、そのカットポイントはフレームの色ヒストグラム

から決定される。互いに隣接するフレームの色ヒストグラム間の差が所定の閾値より大きくなったとき1つのカットポイントが決定される。しかしながら、この方法は、キーフレームを多く生成し過ぎ、多くの場合（例えば、映画、ニュース、レポートなど）で、選択されたキーフレームは多くの同様なフレーム（例えば、新聞のフレーム）を含むことがある。

【0007】このようなヒストグラム技法は、画素ベースであるか、或いはブロックベースである。そして、閾値を決定する方法がシーン変化を判定するのに利用される。このような技法では、照明の変化が同じシーンを示す連続するフレーム間で色をシフトさせるので、誤りの結果をしばしば生成している。同様に、カメラズームショットが多過ぎるキーフレームをしばしば生成している。

【0008】Ratakonda等の米国特許第5,995,095号は、ビデオシーケンスのキーフレームに基づいて階層的なサマリを生成すること、及びデジタルビデオシーケンスのデジタルビデオ信号を入力することを含む、階層的デジタルビデオ要約化及びブラウジングの方法を記載している。この階層的なサマリは、詳細（即ち、フレームの数）によって変化する複数のレベルを含む。最も粗い、即ち、最もコンパクトなレベルでは最も顕著な特徴が形成され、このレベルは最も少ない数のフレームを含む。

【0009】ユーザは、最もコンパクトな（最も粗い）レベルのサマリ、即ち、最もコンパクトなサマリを得ることができる。そして、ユーザは親フレームにタグ付けし、より密なレベルの子（複数の子）フレームを見ることができる。最も密なレベルにおけるフレームにタグ付けすることはビデオの再生に帰着する。その方法は、ショット境界検出を利用して階層の最も密なレベルに含まれるもののためにキーフレームを選択する。ショット境界検出は、ショット境界（即ち、シーン変化）を判定するために互いに連続するフレームのヒストグラム間の差が比較される、閾値法を用いて行われる。この階層的ビデオ要約化方法は、ビットストリームの最小限の復号化と共にMPEG圧縮されたビデオに対して行われる。このビデオ要約化方法は、任意のパン運動の画像モザイク及び任意のズームのズームサマリを任意選択でかつ別々に判定することができる。しかしながら、Ratakonda等は、自動パン／ズーム検出／抽出機能を組み込むには、フレームビットストリーム全体が復号されなければならないことを開示している。更に、Ratakonda等は、計算にコストがかかり計算が非効率的である画素レベルでの動きベクトルに基づくパン及びズーム検出方法を開示している。また、Ratakonda等は、背景／前景の変化又は複雑なカメラ効果が現れる可能性がある実世界の複雑なショットでは効果的に実現することのできないショットフレームのパノラマビューの画像モザイクを構築することも記載している。

10

20

30

40

50

## 【0010】

【発明が解決しようとする課題】本発明は、既存の構成の1つ又は複数の欠点を実質的に解消する、換言すれば少なくとも改善することを目的とする。

## 【0011】

【課題を解決するための手段】本発明の一態様によれば、動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスからキーフレームを抽出する方法であって、前記動きベクトルに基づいてグローバル動き信号を生成する工程と、生成されたグローバル動き信号に基づいてドミナントグローバル方向クラスタを生成する工程と、生成されたドミナントグローバル方向クラスタを使用してキーフレームを選択する工程と、キーフレームを抽出するために前記選択されたキーフレームを解凍する工程とを有することを特徴とする方法が提供される。

【0012】本発明の他の態様によれば、ブロック動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスの表現を生成する方法であって、前記ブロック動きベクトルを得るために前記圧縮ビデオデータを解凍する工程と、前記ブロック動きベクトルを順方向ブロック動きベクトルに変換する工程と、前記順方向ブロック動きベクトルに基づいてグローバル動き信号を生成する工程と、生成されたグローバル動き信号に基づいてドミナントグローバル方向クラスタを生成する工程と、生成されたドミナントグローバル方向クラスタ及び定義済み規則のセットを使用してビデオシーケンスの潜在的なキーフレームを選択する工程と、選択された潜在的なキーフレームの冗長なキーフレームを削除して残りの選択されたキーフレームを得る工程と、前記ビデオシーケンスの表現を得るために前記残りの選択されたキーフレームを解凍する工程とを有することを特徴とする方法が提供される。

【0013】本発明の他の態様によれば、各ビデオクリップがブロック動きベクトルを持つMPEG圧縮ビデオデータを有する、1つ又は複数のビデオクリップからキーフレームを抽出する方法であって、前記ブロック動きベクトルを得るためにMPEG圧縮ビデオデータを部分的に解凍する工程と、前記ブロック動きベクトルを順方向ブロック動きベクトルに変換する工程と、前記順方向ブロック動きベクトルに基づいてパングローバル動き信号、ズームグローバル動き信号、チルトグローバル動き信号を生成する工程と、生成されたパングローバル動き信号、チルトグローバル動き信号、ズームグローバル動き信号に基づいて左パン動きクラスタ、右パン動きクラスタ、上チルト動きクラスタ、下チルト動きクラスタ、ズームイン動きクラスタ、ズームアウト動きクラスタ、グローバル静止動きクラスタの1つ又は複数を含むドミナントグローバル方向クラスタを生成する工程と、生成されたドミナントグローバル方向クラスタ及び定義済み規則のセットを使用して各ビデオクリップの潜在的なキ

ーフレームを選択する工程と、定義済み発見的規則を使用して前記選択された潜在的なキーフレームから冗長なキーフレームを削除して残りの選択されたキーフレームの第1のセットを得る工程と、色ヒストグラム技法を使用して前記残りの選択されたキーフレームの第1のセットから類似したキーフレーム及び／又は繰り返されるキーフレームを削除して残りの選択されたキーフレームの第2のセットを得る工程と、キーフレームを抽出するために前記残りの選択されたキーフレームの第2のセットを解凍する工程とを有することを特徴とする方法が提供される。

【0014】本発明の他の態様によれば、動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスからキーフレームを抽出するコンピュータプログラムを有するコンピュータ読み取り可能な媒体であって、前記コンピュータプログラムは、前記動きベクトルに基づいてグローバル動き信号を生成するコードと、生成されたグローバル動き信号に基づいてドミナントグローバル方向クラスタを生成するコードと、生成されたドミナントグローバル方向クラスタを使用してキーフレームを選択するコードと、キーフレームを抽出するために前記選択されたキーフレームを解凍するコードとを有することを特徴とするコンピュータ読み取り可能な媒体が提供される。

【0015】本発明の他の態様によれば、ブロック動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスの表現を生成するコンピュータプログラムを有するコンピュータ読み取り可能な媒体であって、前記コンピュータプログラムは、前記ブロック動きベクトルを得るために前記圧縮ビデオデータを解凍するコードと、前記ブロック動きベクトルを順方向ブロック動きベクトルに変換するコードと、前記順方向ブロック動きベクトルに基づいてグローバル動き信号を生成するコードと、生成されたグローバル動き信号に基づいてドミナントグローバル方向クラスタを生成するコードと、生成されたドミナントグローバル方向クラスタ及び定義済み規則のセットを使用してビデオシーケンスの潜在的なキーフレームを選択するコードと、選択された潜在的なキーフレームの冗長なキーフレームを削除して残りの選択されたキーフレームを得るコードと、前記ビデオシーケンスの表現を得るために前記残りの選択されたキーフレームを解凍するコードとを有することを特徴とするコンピュータ読み取り可能な媒体が提供される。

【0016】本発明の他の態様によれば、各ビデオクリップがブロック動きベクトルを持つMPEG圧縮ビデオデータを有する、1つ又は複数のビデオクリップからキーフレームを抽出するコンピュータプログラムを有するコンピュータ読み取り可能な媒体であって、前記コンピュータプログラムは、前記ブロック動きベクトルを得るためにMPEG圧縮ビデオデータを部分的に解凍するコードと、前記ブロック動きベクトルを順方向ブロック動

10

20

30

40

50



きベクトルに変換するコードと、前記順方向ブロック動きベクトルに基づいてパングローバル動き信号、ズームグローバル動き信号、チルトグローバル動き信号を生成するコードと、生成されたパングローバル動き信号、チルトグローバル動き信号、ズームグローバル動き信号に基づいて左パン動きクラスタ、右パン動きクラスタ、上チルト動きクラスタ、下チルト動きクラスタ、ズームイン動きクラスタ、ズームアウト動きクラスタ、グローバル静止動きクラスタの1つ又は複数を含むドミナントグローバル方向クラスタを生成するコードと、生成されたドミナントグローバル方向クラスタ及び定義済み規則のセットを使用して各ビデオクリップの潜在的なキーフレームを選択するコードと、定義済み発見的規則を使用して前記選択された潜在的なキーフレームから冗長なキーフレームを削除して残りの選択されたキーフレームの第1のセットを得るコードと、色ヒストグラム技法を使用して前記残りの選択されたキーフレームの第1のセットから類似したキーフレーム及び／又は繰返されるキーフレームを削除して残りの選択されたキーフレームの第2のセットを得るコードと、キーフレームを抽出するために前記残りの選択されたキーフレームの第2のセットを解凍するコードとを有することを特徴とするコンピュータ読み取り可能な媒体が提供される。

【0017】本発明の他の態様によれば、前記何れか1つの方法で作成されるビデオサマリが提供される。

【0018】本発明の他の態様によれば、動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスからキーフレームを抽出する方法であって、前記動きベクトルに基づいてグローバル動き信号を生成する工程と、生成されたグローバル動き信号に基づいてキーフレームを選択する工程と、前記選択されたキーフレームを解凍する工程とを有することを特徴とする方法が提供される。

【0019】本発明の他の態様によれば、動きベクトルを持つ圧縮ビデオデータを有するビデオシーケンスからキーフレームを抽出する装置であって、前記動きベクトルに基づいてグローバル動き信号を生成する手段と、生成されたグローバル動き信号に基づいてキーフレームを選択する手段と、前記選択されたキーフレームを解凍する手段とを有することを特徴とする装置が提供される。

【0020】

【発明の実施の形態】以下、図面を参照しながら本発明のいくつかの好ましい実施形態について説明する。

【0021】図面のうち任意の1つ又は複数において、参照は同じ参照符号の工程及び／又は特徴に対してなされ、これらのステップ及び／又は特徴は矛盾する趣旨が現れない限り、説明の都合上、同じ機能を有し又は動作する。

【0022】ビデオカメラのユーザは、しばしば1つの位置から他の位置にパン及びズームして異なった場所及び事象の関連性を示し、カメラを静止状態に保持し、重

要な事象又はユーザが関心を有する特定のものに焦点を合わせている。この実施形態によるキーフレーム抽出方法は、圧縮されたビデオデータから推定されるカメラ運動のドミナントグローバル方向クラスタに基づかれるものである。この方法は、入力ビデオが既に圧縮された形態であるという事実を利用している。この利点は、あらゆるフレームを完全に解凍する際の計算上のコストが不要になることである。選択された多数のキーフレームをプロセスの終了時に解凍するだけでよい。また、この方法はユーザの関心といくつかの重要な事象を取り込むことを試みている。この方法は、ビデオの複雑さによる合理的な多数の効率的かつ効果的なキーフレームを与えている。

【0023】ここで説明する好ましい方法の原理は、ビデオシーケンスからキーフレームを抽出する方法に一般的に適用可能である。しかしながら、説明の都合上、ビデオクリップを参照しながら好ましい方法のステップを説明する。ビデオクリップは、記録開始取込み事象及び記録終了取込み事象との間のビデオの区分として定義される。しかしながら、本発明は、後述する方法には限らない。例えば、本発明は多数のそのようなクリップを有する商業的な映画などに適用しても良い。また、この方法はMPEG (Moving Picture Coding Experts Group) 2などの動き補償予測圧縮ビデオに適用可能である。しかしながら、この方法はこれに限らない。動きベクトルを組み込んだあらゆる圧縮ビデオシーケンスに適切である。

【0024】次に、図1を参照して、第1の実施形態によるビデオクリップからキーフレームを抽出する方法の概要に関する流れ図を示す。キーフレーム抽出方法100はステップ102から開始され、任意の必要なパラメータが初期設定される。方法100は続くステップ104で、方法100によって処理するMPEG2の圧縮ビデオクリップを入力する。そして、方法100はステップ106に進み、ビデオクリップのグローバル動き信号を生成する。これらのグローバル動き信号は、ビデオクリップの大部分のフレームのためのグローバル動きパラメータを備えている。これらのグローバル動きパラメータは、ビデオクリップの利用可能な各フレーム毎にパンパラメータ、チルトパラメータ、ズームパラメータを備えている。グローバル動き信号は、ビデオシーケンス全体について時間の関数としてこれらのパラメータを備えている。例えば、図5Aから図5Cは、例示的なビデオシーケンスに関するグローバル動き信号を示すグラフである。これらのグローバル動き信号を生成する方法については以下に詳しく説明する。

【0025】ステップ106の後、方法はステップ108に進み、ビデオシーケンス全体に関するグローバル動きのドミナント方向のクラスタ（以下、「ドミナントグローバル方向クラスタ」と呼ぶ）が生成される。この方

10

20

30

40

50



法では、このステップ108の間、利用可能な各フレーム毎に全3つのグローバル動き信号を入力として取り込み、フレームのグローバル動き信号のドミナント方向を判定する。フレームのドミナント方向は、左パン、右パン、上チルト、下チルト、ズームイン、ズームアウト、グローバル静止の何れかで良い。そして、全ての利用可能なフレームのドミナント方向をクラスタ化し、ビデオシーケンス全体のドミナントグローバル方向クラスタを形成する。例えば、図5Dは、図5Aに用いられたのと同じ例示的なビデオシーケンスの時間の関数としてのドミナントグローバル方向クラスタ（即ち、フレーム番号）のグラフである。図5Dを見るとわかるように、ビデオシーケンスは、グローバル静止クラスタを有するフレーム番号=0から始まり、ズームインクラスタ、左パンクラスタに進み、グローバル静止クラスタで終わるクラスタにセグメント化される。これらのドミナントグローバル方向クラスタがどのように生成されるかについては以下に詳しく説明する。

【0026】ステップ108の後、方法はステップ110に進み、ビデオシーケンスからキーフレームを抽出する。このステップ110の間、1組の予め定義された規則のセットを用いてドミナントグローバル方向クラスタからいくつかの潜在的なキーフレームが選択される。そして、この方法は冗長な潜在的なキーフレームを削除して、最後に結果として得られる残りのキーフレームを完全に復号する。これらのキーフレームがどのように抽出されるかについては以下に詳しく説明する。そして、この方法は復号されたキーフレームをビデオシーケンスのサマリとして出力する。

【0027】次に、図2を参照して、図1のキーフレーム抽出方法のステップ106のサブステップの流れ図を詳細に示す。ステップ104でMPEGビデオシーケンスが入力された後、方法はステップ202に進む。このステップ202の間、MPEG2圧縮ビデオシーケンスが既知の方式で部分的に解凍され、ビデオシーケンスの全てのMPEG2ブロック動きベクトルが得られる。

【0028】動画に関するMPEG2圧縮基準は、ビデオシーケンスの空間的冗長性と時間的冗長性の両方を活用するものである。MPEG2は圧縮のいくつかのモードを利用している。1つのモードはフレーム内符号化（イントラ符号化）と呼ばれ、ビデオの多数の画面（pictures）は個別にかつ独立に圧縮又は符号化される。フレーム内符号化は、画面の互いに隣接する画素間に存在する空間的冗長性を活用している。フレーム内符号化のみを用いて符号化された画面をIピクチャと呼ぶ。MPEGは画面間の時間的冗長性を活用する、フレーム間符号化（インター符号化）と呼ばれる別のモードを利用している。時間的冗長性は、互いに隣接する画面間の高い相関の度合に起因する。

【0029】MPEG2は、予測誤差と呼ばれるフレ

ム間差分信号を算出することによりこの冗長性を活用している。MPEG2は、予測誤差を算出する際に、動き補償に対してマクロブロック手法を使用している。符号化すべきフレームの中の目標マクロブロックが参照画像と呼ばれる前（又は連続）フレームの中の最も類似した置換されたマクロブロックとマッチされる。目標マクロブロックから予測マクロブロックへの置換を記述する

（ブロック）動きベクトルは、最もマッチするマクロブロック、即ち、予測マクロブロックの位置を示す。その（ブロック）動きベクトル情報は符号化され、圧縮画像フレームと共に伝送される。順方向予測では、符号化すべき画面の中の目標マクロブロックが、参照画面と呼ばれる過去の画面の中の同じサイズの1組の置換されたマクロブロックとマッチされる。目標マクロブロックから予測マクロブロックへの水平及び垂直置換を記述する

（ブロック）動きベクトルは、この最もマッチする予測マクロブロックの位置を示す。順方向予測を使用してMPEG2で符号化された画面をPピクチャと呼ぶ。

【0030】また、MPEG2圧縮基準は双方向時間予測を使用している。双方向予測で符号化される画面は2つの参照画面、過去の参照画面、将来の参照画面を使用している。双方向に符号化された画面の中の目標マクロブロックは、過去の参照画面からの予測マクロブロックによって予測するか（順方向予測）、将来の参照画面からの予測マクロブロックによって予測するか（逆方向予測）、2つの予測マクロブロックの平均によって予測するか、各参照画面（補間）からの予測マクロブロックによって予測される。すべての場合において、参照画面からの予測マクロブロックは動きベクトルに関連付けられ、その結果、1マクロブロック当たり、最大2つの動きベクトルを双方向予測と共に使用することができる。

【0031】このステップ202の間、キーフレーム抽出方法はビデオシーケンスを解凍し、ビデオシーケンスのすべてのMPEGブロック動きベクトルを得る。この方法はシーケンスを完全に復号することなく、このステップの間にフレーム間復号が行われることはない。当業者には、すべてのフレームが部分的に復号されるわけではないことが理解されよう。即ち、すべてのフレームが動きブロックベクトルを有するわけではない。

【0032】この方法は、ステップ202でMPEGビデオシーケンスを部分的に解凍した後、ステップ204に進む。ステップ204では、すべての逆方向（ブロック）動きベクトルが、基本的に参照及び方向の変化を必要とする順方向（ブロック）動きベクトルに変換される。これらの順方向（ブロック）動きベクトルは、グローバル動きを算出することのできるローカル変位ベクトルを表す。

【0033】次に、この方法はステップ206に続き、順方向動きベクトルを含む利用可能な各フレーム毎の3つのグローバル動きパラメータを算出する。これらのグ

10

20

30

40

50

ローカル動きパラメータは、「Global Zoom/Pan estimation and Compensation for video Compression」(Proc ICASSP91, Yi Tong Tse and Richard Baker著、2725ページから2728ページ)に記載された方法を使用してフレームの順方向動きベクトルから算出される。3つのグローバルパラメータ、即ちパンパラメータとしての $x$ 、チルトパラメータとしての $y$ 、ズームパラメータとしての $z$ が算出される。利用可能な各フレーム毎にグローバル動きパラメータが算出される。そして、ビデオシーケンスに関する時間の関数として、これらのグローバルパラメータから3つのグローバル動き信号、 $X=x(t)$ 、 $Y=y(t)$ 、及び $Z=z(t)$ が形成される。このように生成されたグローバル動き信号の例が図5Aから図5Cに示されている。

【0034】次に、図5Aを参照して、例示的なビデオシーケンスに関する(パン)生成されたグローバル動き信号 $x(t)$ のグラフを示す。水平軸はフレーム番号0からフレーム番号260までのビデオシーケンスのフレームの数を表している。垂直軸は-10画素から+20画素まで変化するパンパラメータ $x$ である。パンパラメータ $x$ が-10から+10までは一般に小さなパン移動を示すことがわかる。10を越えるパンパラメータ $x$ は左パン移動を示す。-10未満のパンパラメータ $x$ は右パン移動を示す。図からわかるように、ビデオシーケンスの間中、概ねフレーム番号130から230までは左パン移動である。

【0035】次に、図5Bを参照して、図5Aで使用されているのと同じ例示的なビデオシーケンスの(チルト)生成されたグローバル動き信号 $y(t)$ のグラフを示す。図5Aと同様に、水平軸はフレーム番号0からフレーム番号260までのビデオシーケンスのフレームの数を表している。垂直軸は-20画素から+20画素まで変化するチルトパラメータ $y$ である。チルトパラメータ $y$ が-10から+10までは一般に小さなチルト移動を示すことがわかる。10を越えるチルトパラメータ $y$ は上チルト移動を示す。-10未満のパンパラメータ $y$ は下チルト移動を示す。図からわかるように、ビデオシーケンスの間中、概ねフレーム番号60、125、150、160、220、及び240は短いチルト移動である。

【0036】次に、図5Cを参照して、図5Aで使用されているのと同じ例示的なビデオシーケンスの(ズーム)生成されたグローバル動き信号 $z(t)$ のグラフを示す。図5Aと同様に、水平軸はフレーム番号0からフレーム番号260までのビデオシーケンスのフレームの数を表している。垂直軸は0ズーム係数から1.2ズーム係数まで変化するズームパラメータ $z$ である。ズームパラメータ $z$ が0.98から1.02までは一般に小さなズーム移動を示すことがわかる。1.02を超えるズームパラメータ $z$ はズームアウト移動し、0.98未満

のズームパラメータ $z$ はズームイン移動を示す。図からわかるように、ビデオシーケンスの間中、概ねフレーム番号40から140まではズームイン移動である。

【0037】次に、図3を参照して、図1に示した方法のステップ108のサブステップの流れ図を詳細に示す。キーフレーム抽出方法はステップ206でグローバル動き信号を算出した後、ステップ302に進む。ステップ302の間、方法はグローバル信号の各々の閾値を導入する。即ち、キーフレーム抽出方法は、利用可能な各フレームの各グローバルパラメータを3つの離散グローバルパラメータ値+1、0、-1の1つに変換する。例えば、グローバルパンパラメータ $x$ の場合に、 $-10 \leq x \leq +10$ であれば $x$ が0に変換される。さもなければ、 $x > 10$ であれば $x$ が1に変換され、 $x < -10$ であれば $x$ が-1に変換される。グローバルチルトパラメータ $y$ についても同様である。グローバルズームパラメータ $z$ の場合に、 $0.98 \leq z \leq 1.02$ であれば $z$ が0に変換される。さもなければ、 $z < 0.98$ であれば $z$ が-1に変換され、 $z > 1.02$ であれば $z$ が+1に変換される。

【0038】閾値を導入するステップ302の後、ステップ304に進み、離散グローバル動き信号からノイズを除去する。離散グローバルパラメータがより顕著な形状に変形される既知の形態論処理技法を使用して離散グローバル動き信号からノイズを除去する。この技法は、短い遷移スパイクを除去し、離散グローバル動き信号中のあらゆるホールを充填する。形態論処理の基本動作を説明した「Fundamentals of Digital Image Processing」(A.K.Jain著、384ページ)を参照されたい。この方法は、ステップ304の後、ステップ306に進む。

【0039】このキーフレーム抽出方法は、ステップ306の間、ビデオシーケンス全体に渡ってノイズが低減された離散グローバル動き信号に基づいてドミナントグローバル方向クラスタを生成する。この方法は、利用可能な各フレーム毎のノイズが低減されたすべての3つの離散グローバル動き信号を入力として取り込み、このフレームのグローバル動き信号のドミナント方向を判定する。フレームのドミナント方向は、左パン、右パン、上チルト、下チルト、ズームイン、ズームアウト、グローバル静止の何れかである。そして、すべての利用可能なフレームのドミナント方向をクラスタ化し、ビデオシーケンス全体のドミナントグローバル方向クラスタを形成する。例えば、「グローバル静止」運動はカメラ静止及び/又は小さなローカルオブジェクト運動、即ちゼロに近い又はゼロに等しいそのフレームに対するすべての離散グローバル動き値を取り込む。左パン運動の例では、フレームの離散グローバル動き値は(パン=1、ズーム=0、チルト=0)である。しかしながら、フレームに関して組み合わせられた動きが存在する場合(例えば、離散グローバル信号パン=1、ズーム=1、チルト

=0)、グローバル運動のドミナント方向は、ステップ206で算出された3つの最初のグローバル動き信号のうちの最大の信号である。後者の場合、比較が行われ、フレーム中の最大の最初のグローバル動き信号を判定する。好ましくは、まず比較の前に、時間の経過と共に最初のグローバル動き信号の平均を求め、あらゆる遷移を削除する。図5Aから図5Cに示されるように、ズームグローバルパラメータは、チルトグローバルパラメータやパングローバルパラメータとは異なるメトリックを有する。ズームグローバル動き信号は、チルトグローバル動きパラメータ又はパングローバル動きパラメータと直接比較してグローバル運動のドミナント方向を判定できるように正規化される。各フレーム毎のグローバル運動のドミナント方向を判定した後、これらのドミナント方向をグループ化してクラスタを形成する。前述のように、このようなクラスタを本実施形態ではドミナントグローバル方向クラスタと呼ぶ。このクラスタ化サブステップにより、同じ種類の連続的で密なクラスタがグループ化される。非常に短い動きセグメントは無視される。

【0040】次に、図5Dを参照して、図5Aで使用されているのと同じ例示的なビデオシーケンスの時間の関数としてのドミナントグローバル方向クラスタ（即ち、フレーム番号）のグラフを示す。図示するように、フレーム番号0から40までのグローバル静止クラスタがある。この時間中のドミナントグローバル運動はなく、ビデオシーケンスは比較的静止している。フレーム番号40から130の間にズームインクラスタがある。この時間中のドミナントグローバル運動はズームインである。ズームインクラスタの後、フレーム番号130から230の間に左パンクラスタがある。この持続時間の間のドミナントグローバル運動は左パンである。最後のクラスタは、フレーム番号230から260までのグローバル静止クラスタである。この期間中のドミナントグローバル運動はない。

【0041】次に、図5Aから図5Cに示したように、フレーム番号130から140までの期間に左パン移動とズームイン移動とが重なり合っていることがわかる。このキーフレーム抽出方法は、この期間中ドミナントグローバル運動を左パンであると判定する。概ねフレーム番号60、125、150、160、220でのチルト移動が左パン移動とズームイン移動の両方と重なり合っていることもわかる。しかしながら、チルト動きセグメントは短く、キーフレーム抽出方法はクラスタ化の間このようなセグメントを無視している。

【0042】次に、図4を参照して、図1のキーフレーム抽出方法のステップ110のサブステップの流れ図を示す。ステップ306でドミナントグローバル方向クラスタを判定した後、この方法はステップ402に進む。このステップ402の間、ドミナントグローバル方向クラスタから潜在的なキーフレームが選択される。これら

キーフレームは、この段階でMPEG圧縮ビデオから抽出されることはない。これらキーフレームは、以下の定義済み規則のセットに応じて選択される。

【0043】・パン又はチルトの長さ及び速度によってパンクラスタ又はチルトクラスタのために1つ又は複数のフレームを選択する。

【0044】・ズーム係数及びズームの長さによってズームクラスタのために1つ又は複数のフレームを選択する。

【0045】・各グローバル静止クラスタのために1つのフレームのみを選択する。

【0046】例えば、パンクラスタから選択される潜在的なキーフレームは、パンクラスタの長さ及び速度によってパンの開始部分、パンの中央、及びパン終了部分に存在することも、或いはパンの開始部分から1つの潜在的なキーフレームにのみ存在することもできる。

【0047】次に、図5Dに戻り、図5Aから図5Cに示されるビデオシーケンスのために選択された潜在的なキーフレームを説明する。この例では、第1のグローバル静止クラスタのために選択される1つのキーフレーム、ズームインクラスタの開始及び終了でズームインクラスタのために選択される2つのキーフレーム、左パンクラスタの開始及び終了で選択される2つのキーフレーム、最後のグローバル静止クラスタのために選択される1つのキーフレームがある。

【0048】ステップ402の後、この方法はステップ404に進む。ステップ404では、ステップ402で選択されたキーフレームの冗長なキーフレームが写真の特性を考慮した定義済みの発見的規則のセットに基づいて削除される。例えば、グローバル静止は他の動きクラスタよりも重要である。グローバル静止では、しばしばユーザの関心又はある特定の事象が取り込まれる。また、グローバル静止では時には動く画像よりも多くの焦点が合った画像が取り込まれる。適用できる発見的規則のうちのいくつかを以下に示す。

【0049】・「グローバル静止」でないクラスタの開始/終了で選択される潜在的なキーフレームが「グローバル静止」クラスタの後に続くか、このキーフレームの後に「グローバル静止」クラスタが続く場合、このキーフレームを削除する。例えば、「ズームイン」クラスタの終了で選択される潜在的なキーフレームの後に「グローバル静止」クラスタが続くとき、このキーフレームを削除する。「グローバル静止」から選択されるキーフレームは、しばしば「ズームイン」クラスタの終了から抽出されるキーフレームよりも良い品質を有し、これらのキーフレームは類似している。

【0050】・パン/チルトの開始で選択される潜在的なキーフレームがズームクラスタの後に続く場合、このキーフレームを削除する。例えば、「ズームイン」クラスタの後に、非常に短い無視される動きセグメントが続

き、その後、「右パン」クラスタが続く場合は、「右パン」の開始からの潜在的なキーフレームを削除する。しかしながら、発見的規則のセットはこれらの使用規則に限らない。

【0051】そして、この方法はステップ406に進み、ステップ404の後、残っている選択されたキーフレームのうち同様なキーフレーム及び／又は繰り返されるキーフレームを削除する。定義済みの発見的規則を使用するにもかかわらず、依然としてシーンが繰り返され、異なる時間に同様なキーフレームが生じることもある。このステップ404は、画像類似性測定を用いてこれらの同様なキーフレームを削除する。画像の類似性を測定する既存の方法（例えば、色ヒストグラム比較）が使用できる。色ヒストグラムを算出する際に、PフレームやBフレームではなく、MPEGビデオの最も密なIフレームがキーフレームとして使用される。その場合、MPEG圧縮画像のDC係数を使用して低解像度画像を生成することができる。画像の類似性の測定は、DCキーフレーム画像を使用して行うことができる。

【0052】ステップ406の後、この方法はステップ408に進み、ステップ406から依然として残っている選択されたキーフレームをMPEGビデオシーケンスから完全に復号する。これらの復号された残りキーフレームは、MPEGビデオシーケンスのサマリを形成し、ステップ112で出力される（例えば、ハードディスク上に記憶される）。そして、この方法はステップ114で終了する。このように、この方法は高速で効率的なキーフレーム抽出を実行する。

【0053】次に、図5Dに戻り、図5Aから図5Cに示されるビデオシーケンスのために抽出されたキーフレームを説明する。ステップ404の間、図示するように抽出されたキーフレームを残す上述の発見的規則のセットを用いて第2、第4、及び第5の潜在的なキーフレームが除去される。この例における削除ステップ406は、同様なキーフレームも繰り返されるキーフレームも見つからず、抽出されたキーフレームは図示のように残される。そして、この方法はこれらの抽出されたキーフレームを復号し、ビデオシーケンスのビデオサマリを形成する。

【0054】【装置の好ましい実施形態】キーフレームを抽出する方法は、図6に示すような従来型の汎用コンピュータシステム600を使用して実施することが好ましい。ここで、図6のプロセスはコンピュータシステム600の中で実行するアプリケーションプログラムなどのソフトウェアとして実施される。特に、キーフレームを抽出する方法のステップは、コンピュータによって実行されるソフトウェアとして符号化された命令によって実施される。このソフトウェアは、キーフレーム抽出方法を実行する1つの部分と、この部分とユーザとの間のユーザインタフェースを管理する他の部分との2つの別

々の部分に分割することができる。このソフトウェアは、例えば後述する記憶装置を含むコンピュータ読み取り可能な媒体に記憶されても良い。このソフトウェアは、コンピュータ読み取り可能な媒体からコンピュータにロードされ、そしてコンピュータによって実行される。このようなソフトウェア又はコンピュータプログラムが記録されたコンピュータ読み取り可能な媒体はコンピュータプログラム製品である。コンピュータ内でこのコンピュータプログラム製品を使用することにより、本発明の実施形態に応じてキーフレームを抽出する有利な装置を実現することが好ましい。

【0055】コンピュータシステム600は、コンピュータモジュール601、キーボード602及びマウス603のような入力装置、プリンタ615を含む出力装置及び表示装置614を備えている。変復調（モデム）トランシーバ装置616が、例えば電話回線621又はその他の機能媒体を介して接続可能な通信ネットワーク620との間で通信するためにコンピュータモジュール601によって使用される。このモデム616は、インターネット、及びローカルエリアネットワーク（LAN）やワイドエリアネットワーク（WAN）など他のネットワークシステムへのアクセスを得るために使用される。

【0056】コンピュータモジュール601は通常、少なくとも1つのプロセッサユニット605、例えば半導体ランダムアクセスメモリ（RAM）や読み取り専用メモリ（ROM）で形成されたメモリユニット606、ビデオインタフェース607を含む入出力（I/O）インタフェース、キーボード602及びマウス603及びオプションのジョイスティック（図示せず）用の入出力インタフェース613、並びにモデム616用のインタフェース608を含む。記憶装置609が設けられ、この装置は通常、ハードディスクドライブ610及びフロッピーディスクドライブ611を含む。また、磁気テープドライブ（図示せず）を使用することもできる。非揮発性データ源として通常、CD-ROMドライブ又はDVDドライブ612が設けられている。コンピュータモジュール601の構成要素605～613は通常、相互接続されたバス604を介して、関連技術におけるこれら周知のコンピュータシステム600の従来型の動作モードが得られるように通信している。本実施形態を実施できるコンピュータの例には、IBM-PC及びコンパチブル、Sun Sparcstation又はそれらから進化した同様なコンピュータシステムが含まれる。

【0057】通常、好ましい実施形態のアプリケーションプログラムは、ハードディスクドライブ610上に存在し、プロセッサ605によって実行される際に読み取られ制御される。プログラム及びネットワーク620から取り込まれたあらゆるデータの間記憶は、半導体メモリ606を、場合によってはハードディスクドライブ610と共に使用して行われる。いくつかの例では、ア

アプリケーションプログラムは、CD-ROM又はフロッピーディスク上に符号化されてユーザに供給され、対応するドライブ612又は611を介して読み取るか、或いはユーザによってモデム装置616を介してネットワーク620から読み取られても良い。更に、ソフトウェアは磁気テープ、ROM又は集積回路を含む他のコンピュータ読み取り可能な媒体、光磁気ディスク、コンピュータモジュール601及び他の装置の間の無線伝送チャンネル又は赤外線伝送チャンネル、PCMCIAカードなどのコンピュータ読み取り可能なカード、電子メール伝送及びWebサイトなどに記録された情報を含むインターネット及びイントラネットからコンピュータシステム600にロードされても良い。上述したものは、関連するコンピュータ読み取り可能な媒体の例に過ぎない。本発明の範囲及び趣旨から逸脱せずに他のコンピュータ可読媒体を実現することができる。

【0058】このコンピュータシステム600は、キーフレーム抽出方法への入力として働く多量のビデオデータを記憶する機能を有する。ビデオデータはDVD-ROMドライブ612を介して、或いはカムコーダ(図示せず)から入出力インタフェース608を介して直接、コンピュータシステム600に入力されても良い。

【0059】キーフレーム抽出方法は、図1に示す機能又は副機能を実行する1つ又は複数の集積回路などの専用ハードウェアで実現されても良い。このような専用ハードウェアは、カムコーダやVCRなどに組み込まれても良く、グラフィックプロセッサ、デジタル信号プロセッサ又は1つ又は複数のマイクロプロセッサ及び関連するメモリを含んでも良い。

【0060】【産業上の利用可能性】本発明の実施形態がビデオ処理業界に適用可能であることは上述の説明からも明らかである。キーフレーム抽出方法は、例えばビデオ内容の視覚識別、ビデオインデックス付け、ビデオブラウジング、ビデオ編集など多数の用途を有する。ここで、図5Dに戻り、図5Aから図5Cで用いられているビデオシーケンスのための抽出されたキーフレームを示す。ステップ404の間、図示するように、キーフレームを残す上述の発見的規則のセットを使用して第2、第4、及び第5の潜在的なキーフレームが削除される。

【0061】以上、本発明の1つの実施形態/いくつかの実施形態のみについて説明したが、本発明の範囲及び趣旨から逸脱せずに実施形態に修正及び/又は変更を加えることができる。各実施形態は例示的なものであり、制限的なものではない。

#### 【図面の簡単な説明】

【図1】第1の実施形態によってビデオシーケンスからキーフレームを抽出する方法の概要の流れ図である。

【図2】図1に示す方法のステップ106のサブステップの流れ図である。

【図3】図1に示す方法のステップ108のサブステップの流れ図である。

【図4】図1に示す方法のステップ110のサブステップの流れ図である。

【図5A】例示的なビデオシーケンスの(パン)グローバル動き信号 $x(t)$ のグラフである。

【図5B】図5Aと同様に例示的なビデオシーケンスの(チルト)グローバル動き信号 $y(t)$ のグラフである。

【図5C】図5Aと同様に例示的なビデオシーケンスの(ズーム)グローバル動き信号 $z(t)$ のグラフである。

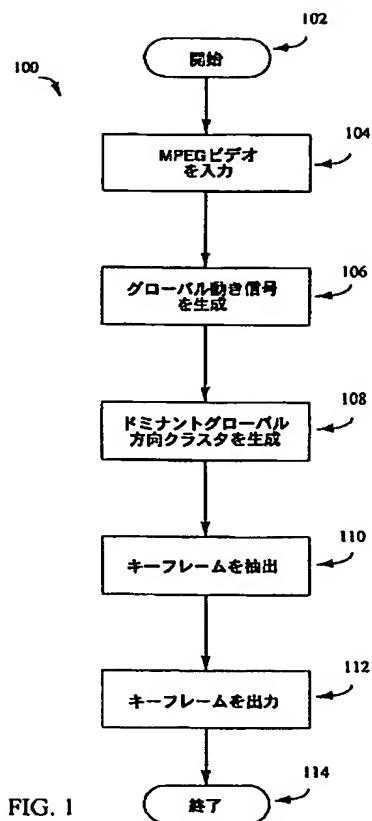
【図5D】図5Aと同様に例示的なビデオシーケンスの、時間の関数としてのドミナントグローバル方向クラスタ、潜在的なキーフレーム、及び生成されるキーフレームのグラフである。

【図6】本発明の実施形態を実施できる汎用コンピュータの概略ブロック図である。

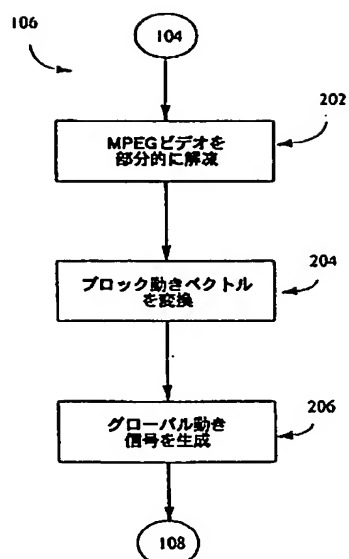
#### 【符号の説明】

- 600 コンピュータシステム
- 601 コンピュータモジュール
- 602 キーボード
- 603 マウス
- 604 相互接続されたバス
- 605 プロセッサユニット
- 606 メモリユニット
- 607 ビデオインタフェース
- 608 入出力(I/O)インタフェース
- 609 記憶装置
- 610 ハードディスクドライブ
- 611 フロッピーディスクドライブ
- 612 CD-ROMドライブ又はDVDドライブ
- 613 入出力インタフェース
- 614 表示装置
- 615 プリンタ
- 616 モデム
- 620 ネットワーク
- 621 電話回線

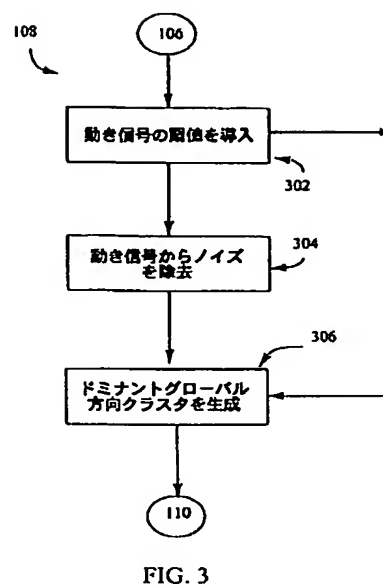
【図1】



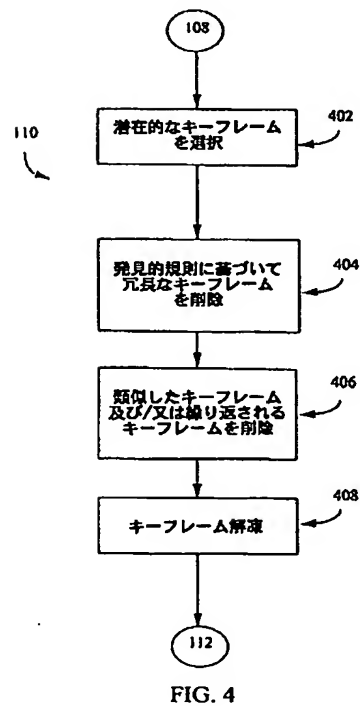
【図2】



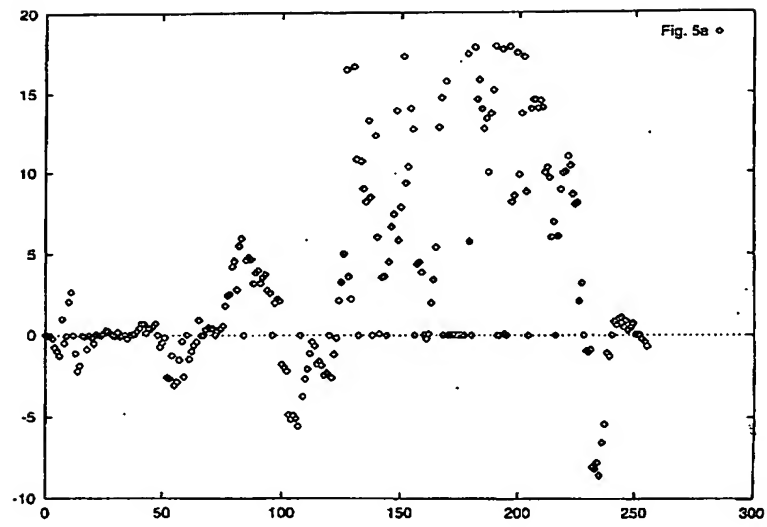
【図3】



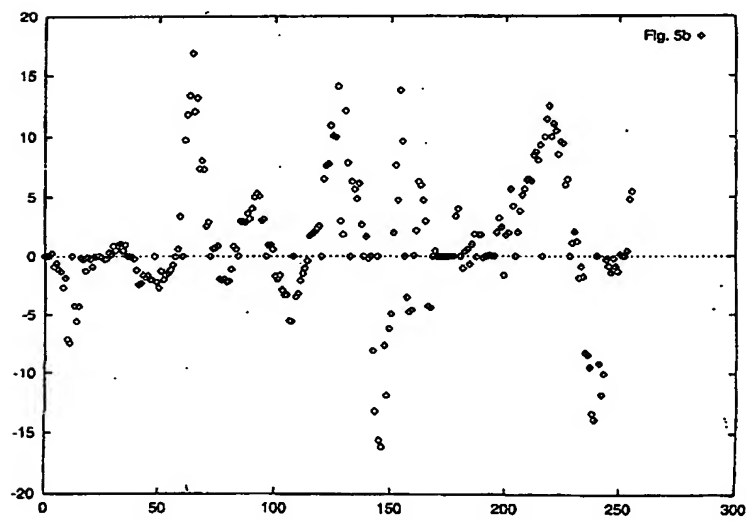
【図4】



【図5A】

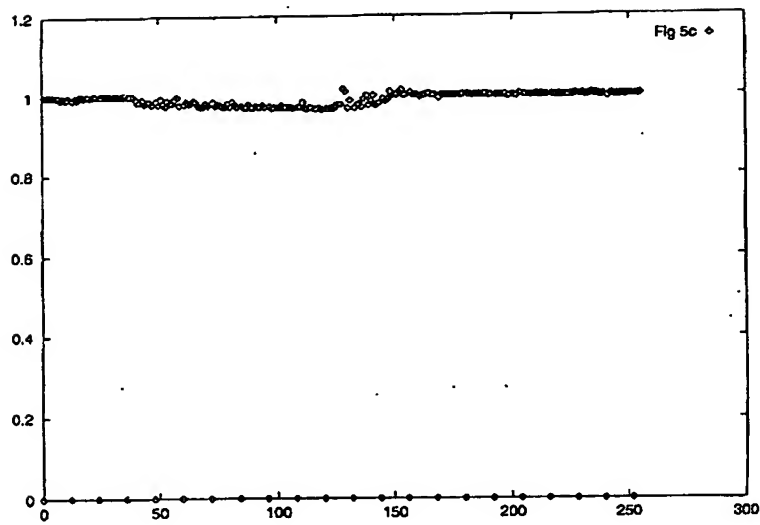


【図5B】





【図5C】



【図5D】

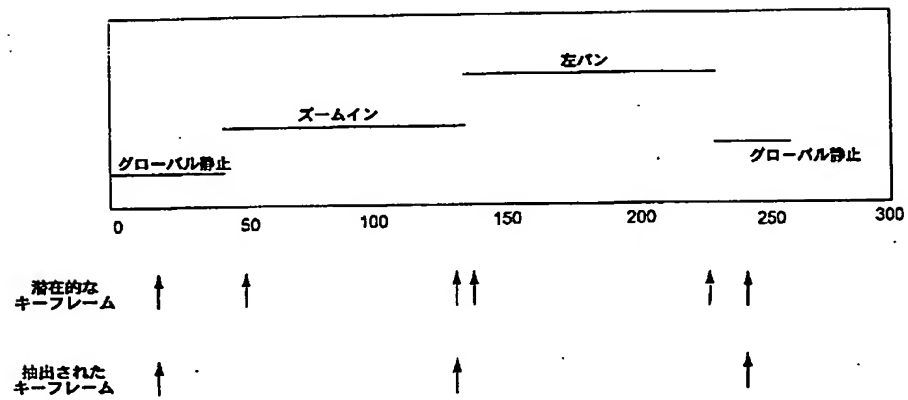


FIG. 5 d

【図6】

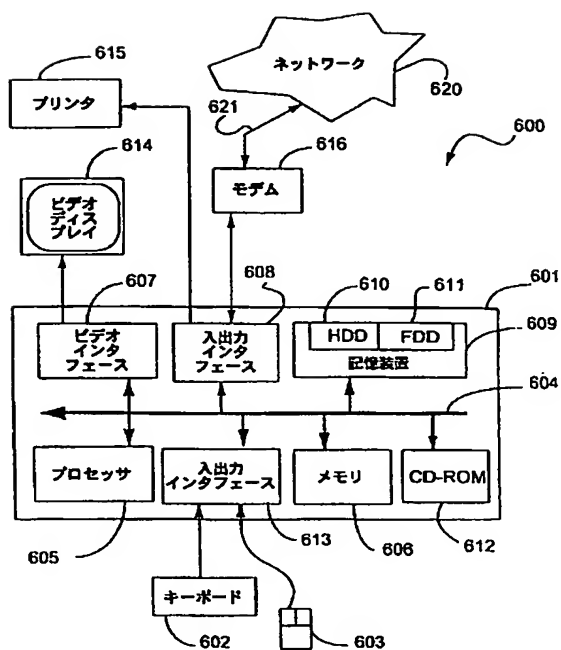


FIG. 6

フロントページの続き

(72)発明者 アリソン ジョアン レノン  
 オーストラリア国 2113 ニュー サウス  
 ウェールズ州、ノース ライド、トーマ  
 ス ホルト ドライブ 1 キヤノン イ  
 ンフォメーション システムズ リサーチ  
 オーストラリア プロプライエタリー  
 リミテッド 内